Biological nanopores having tunable pore diameters and uses thereof as analytical tools

Huang, Kevin; Maglia, Giovanni

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 March 2020 (19.03.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/055246 A1

(51) International Patent Classification:
C07K 14/435 (2006.01) *G01N 33/487* (2006.01)

(21) International Application Number:
PCT/NL2019/050588

(22) International Filing Date:
11 September 2019 (11.09.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
18193722.8 11 September 2018 (11.09.2018) EP

(71) Applicant: **RIJKSUNIVERSITEIT GRONINGEN** [NL/NL]; Broerstraat 5, 9712 CP Groningen (NL).

(72) Inventors: **HUANG, Gang**; University of Groningen, Faculty of Science and Engineering, Groningen Biomolecular Sciences & Biotechnology Institute (GBB), Chemical Biology, Nijenborgh 7, 9747 AG Groningen (NL). **MAGLIA, Giovanni**; University of Groningen, Faculty of Science and Engineering, Groningen Biomolecular Sciences & Biotechnology Institute (GBB), Chemical Biology, Nijenborgh 7, 9747 AG Groningen (NL).

(74) Agent: **WITMANS, H.A.**; V.O., P.O. Box 87930, 2508 DH Den Haag (NL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*

WO 2020/055246 A1

(54) Title: BIOLOGICAL NANOPORES HAVING TUNABLE PORE DIAMETERS AND USES THEREOF AS ANALYTICAL TOOLS

(57) Abstract: The invention relates to the field of nanopores, in particular to engineered Fragaceatoxin C (FraC) nanopores and their application in analyzing biopolymers and other (biological) compounds, such as single-molecule (protein) sequencing. Provided is a system comprising oligomeric FraC nanopores comprised in a lipid bilayer, wherein the sum of the nanopore fraction in the heptameric (Type II) state and the nanopore fraction in the hexameric (Type III) state represents at least 60% of the total number of FraC nanopores.

Title: Biological nanopores having tunable pore diameters and uses thereof as analytical tools.

The invention relates generally to the field of nanopores and the use thereof in analyzing biopolymers and other (biological) compounds. In particular, it relates to engineered Fragaceatoxin C (FraC) nanopores and their application in single molecule analysis, such as single molecule peptide sequencing.

Biological nanopores are proteins that open nanoscale water conduits on biological or synthetic membranes. Under an external potential, the ionic current across single nanopore is used to recognize analytes traversing the nanopore. Most notably, nanopores are now used to sequence nucleic acids at the single-molecule level. In nanopore DNA sequencing, individual nucleic acid trains are threaded base-by-base through nanopores, while the ionic current is used to identify individual nucleobases[1,2].

Watanabe e*t al.* (Analytical Chemistry 2017 89 (21), 11269-11277) describe the analysis of pore formation and detection of a single protein molecule using a large nanopore among five different pore-forming proteins, including FraC. It is demonstrated that the identification of appropriate pores for nanopore sensing can be achieved by classifying the channel current signals and performing noise analysis.

However, the sequencing of proteins with nanopores presents a new set of challenges. Amino acids have a larger chemically variability compared to nucleobases, and they cannot be uniformly captured or stretched by the electrical potential inside the nanopore. Furthermore, enzymes that process proteins or polypeptides amino acid-by-amino acid are not yet known.

Alternatively, proteins might be first fragmented and then the mass of individual peptides identified by nanopore currents. This approach

would be similar to conventional protein sequencing approaches using tandem mass spectrometry. A nanopore peptide mass identifier, however, would have the advantage of being low-cost and portable and single-molecule. The latter is important because it would allow the analysis of the chemical heterogeneity in post-translational modifications and, especially when coupled to high-throughput devices, permit the detection of low-abundance proteins. Previous work with PEG molecules[3,4,5,6,7,8], neutral peptides[9] or oligosaccharides[10], uniformly charged peptides[11,12,13,14,15] and other peptides[16] revealed that there might be a direct correlation between the depth of the current blockade and the molecular weight of polymers, when the composition of the analyte is uniform. On the other hand, a wealth of other studies, including work with DNA[17,18] and amino acid enantiomers[19] revealed that the chemical identity of molecules and especially the charge inside the nanopore[20] have a strong and unpredictable effect on the ionic current, suggesting that the identification of the mass of complex biopolymers such as peptides might not be possible. An additional complication is that peptides of opposite charge are not efficiently captured and analysed at a fixed potential[16,21,22,23,24]. Finally, the diameter and geometry of biological nanopores cannot be easily adapted to study the array of sizes, shapes and chemical composition of polypeptides in solution.

Recently we have shown that octameric Fragaceatoxin C (FraC) nanopores[25] from the sea anemone *Actinia fragacea* can be used to study DNA[26], proteins and peptides[27]. See also WO2018/012963 in the name of the present applicant. The transmembrane region of FraC is unique compared to other nanopores used in biopolymer analysis as it is formed by α-helices that describe a sharp and narrow constriction at the *trans* exit of the nanopore. Crucially, we showed that an electro-osmotic flow across the nanopore can be engineered to capture polypeptides at a fixed potential despite their charge composition[27]. However, peptides smaller than 1.6 kDa in size translocated too fast across the nanopore to be sampled.

Based on these studies, the present inventors realized and recognized that nanopores with a smaller diameter are required to detect peptides with lower molecular weight. Therefore, they aimed at providing a strategy that allows for tuning the diameter of FraC nanopores, such that a larger range of peptides sizes can be identified.

It was surprisingly found that the FraC nanopore can be engineered to induce the formation of different nanopore types (herein referred to Type II and/or Type III) when comprised in the context of a lipid bilayer, thereby creating a biological nanopore with sub-nm constriction. Importantly, these novel, narrow types of nanopores allow for distinguishing (small) peptides differing by the substitution of one amino acid with a ~40 Da resolution, while previous nanopore studies only reported differences of about 200 Da. Moreover, at selected pH conditions the FraC nanopore signal directly correlated to the mass of the peptide. The invention herewith provides a new and unique approach for the single-molecule identification of proteins based on nanopores.

In one embodiment, the invention relates to a system comprising oligomeric Fragaceatoxin C (FraC) nanopores comprised in a lipid bilayer, wherein the sum of the nanopore fraction in the Type II state and the nanopore fraction in the Type III state represents at least 60% of the total number of FraC nanopores.
For example, the sum of the Type II and Type III state nanopores represents at least 65%, preferably at least 70%, of the total number of FraC nanopores.

As used herein, the term "Type II" state refers to nanopores having an apparent heptameric stoichiometry, and/or a conductance of about 1.22-1.26 nS when assayed at pH 7.5 in a 1M NaCl solution or about 0.99-1.08 nS when assayed at pH 4.5 in a 1 M KCl solution. Conductance values are suitably determined by collecting single channels under -50 mV applied

potential using 1 M NaCl, 15 mM Tris pH 7.5, or 1 M KCl, 0.1 M citric acid, 180 mM Tris base pH 4.5.

Type II FraC nanopores are furthermore characterized by an apparent pore size (at the narrowest constriction) of about 1.1 nm as calculated from homology modeling.

As used herein, the term "Type III" state refers to nanopores having an apparent hexameric stoichiometry, and/or a conductance of about 0.37-0.43 nS when assayed at pH 4.5 in a 1M KCl solution.

Type III FraC nanopores are furthermore characterized by a pore size (at the narrowest constriction) of about 0.8 nm as shown by homology modeling.

Accordingly, in one embodiment, the invention provides a system comprising oligomeric FraC nanopores comprised in a lipid bilayer, wherein the sum of the nanopore fraction in the heptameric (Type II) state and the nanopore fraction in the hexameric (Type III) state represents at least 60% of the total number of FraC nanopores.

In another embodiment, the invention provides a system comprising oligomeric FraC nanopores comprised in a lipid bilayer, wherein the sum of (i) the nanopore fraction showing a conductance of about 0.99-1.08 nS (Type II) when assayed at pH 4.5 in a 1 M KCl solution and (ii) the nanopore fraction showing a conductance of about 0.37-0.43 nS (Type III) when assayed at pH 4.5 in a 1 M KCl solution represents at least 60% of the total number of FraC nanopores.

Still further, the invention provides a system comprising oligomeric FraC nanopores comprised in a lipid bilayer, wherein the sum of the nanopore fraction having an apparent pore size of about 1.1 nm (Type II) and the nanopore fraction having an apparent pore size of about 0.8 nM (Type III) represents at least 60% of the total number of FraC nanopores.

The relative amounts of Type II and Type III nanopores can vary according to needs. In one aspect, at least 40%, preferably at least 50%, of the FraC nanopores is in the Type II state. Alternatively, or additionally, at least 15%, preferably at least 20%, of the FraC nanopores is in the Type III state.

In one embodiment, at least 60%, preferably at least 70%, of the FraC nanopores is in the Type II state. In another embodiment, at least 60%, preferably at least 70%, of the FraC nanopores is in the Type III state.

Also encompassed are systems comprising essentially one oligomeric form / Type of FraC. For example, in one embodiment, at least 90%, preferably at least 95%, of the FraC nanopores is present in the Type II state. In a specific aspect, all of the FraC nanopores are in the Type II state. In another embodiment, at least 90%, preferably at least 95%, of the FraC nanopores is present in the Type III state. In a specific aspect, all of the FraC nanopores are in the Type III state. The different oligomeric forms of FraC can be readily isolated using liquid chromatographic techniques, including size-exclusion, affinity, reverse-phase or ion exchange chromatography.

In a specific aspect, the FraC nanopores comprise or consist of mutant FraC monomers comprising one or more mutations that weaken the interaction between the nanopore and the lipid bilayer, i.e. the lipid interface.

Very good results are obtained when FraC is mutated at position W112 and/or W116. For example, in one embodiment, FraC is mutated at position W112, preferably while W116 is not mutated, or at position W116, preferably while W112 is not mutated. In a further embodiment, the FraC mutant comprises a mutation at both positions W112 and W116. According to the present invention, the W residues are substituted with either S, T, A, N, Q or G, preferably with S or T. FraC contains 179 amino acids with relative molecular weight of 20 kDa. The cDNA for FraC is available under the accession number FM958450 in GenBank. The polypeptide sequence of

FraC is available under the accession number B9W5G6 in UniProt. Moreover, the crystal structure of FraC was resolved in 2010 and deposited in the RCSB PDB under the accession number 3LIM. In a mutant according to the invention, the residue numbering corresponds to the residue numbering as in the FraC sequence according to UniProtKB accession number B9W5G6.

The importance of (conserved) tryptophan residues for the functioning of pore-forming toxins has been previously studied. Tanaka *et al.* [25] revealed structures of FraC corresponding to four different stages of its activation route, namely the water-soluble form, the lipid-bound form, an assembly intermediate and the transmembrane pore. Mutational analysis revealed that mutant W112R/W116F lacks the ability to bind to lipid membrane, thus becoming completely inactive. García-Linares *et al.*, 2016 studied the role of the tryptophan residues in the specific interaction of the sea anemone *Stichodactyla helianthus's* Actinoporin Sticholysin II (StnI) with biological membranes. It was found that residues W110 and W114 (corresponding to W112 and W116 of FraC) sustain the hydrophobic effect, which is one of the major driving forces for membrane binding in the presence of cholesterol. Notably, while the authors state that "the results obtained support actinoporins' Trp residues playing a major role in membrane recognition and binding", they also conclude that "their residues have an only minor influence on the diffusion and oligomerization steps needed to assemble a functional pore". Herewith, the findings of the present inventors that W112 and/or W116 are suitably engineered to tune the oligomeric state of FraC could not have reasonably been predicted in view of the prior art.

Hence, in one aspect the invention provides a system comprising oligomeric FraC nanopores comprised in a lipid bilayer, wherein the FraC nanopores comprise mutant FraC monomers comprising a mutation at position W112 and/or W116, preferably wherein the W residue(s) is/ are independently substituted with either S, T, A, N, Q or G, preferably with S or T. In one embodiment, it comprises or consists of FraC mutant W112S or W112T. In

another embodiment, it comprises or consists of mutant W116S or W116T. In a still further embodiment, the system comprises or consists of FraC mutant W112S/W116S, W112T/W116S or W112S/W116T.

The inventors noticed that type II FraC nanopores inserted in the lipid bilayer more efficiently at low pH. Therefore, to increase the production of type II nanopores at physiological pH, aspartic acid at position 109, which is located at the lipid interface, was exchanged for an uncharged residue. Satisfactorily, the fraction of type II nanopores at pH 7.5 increased from 23.0 ± 4.9% to 48 ± 3.6%, and a small fraction of type III nanopores appeared.

The invention also provides a system comprising oligomeric FraC nanopores comprised in a lipid bilayer, wherein the FraC nanopores comprise mutant FraC monomers comprising a mutation at position D109, preferably herein said mutation comprises the substitution of D with an uncharged residue, such as S or T, more preferably with S. The concomitant substitution of tryptophan at position 116 with serine showed a further small increased in the fraction of type I and type II nanopores at pH 7.5

Therefore, the invention further relates to a system comprising oligomeric FraC nanopores comprised in a lipid bilayer, wherein the FraC nanopores comprise mutant FraC monomers comprising mutation D109S and one or both of W112S and W116S.

A system with the nanopores of the invention can accommodate peptides ranging from 22 to 4 amino acids in length. Even smaller peptides can be detected using further fine-tuning of the transmembrane region of the nanopore, for example by introducing amino acids with bulky side-chains. We also showed that the nanopores can discriminate differences between an alanine and glutamate (~40 Da) in mixture of peptides. Furthermore, the inventors found that at exactly pH 3.8 the ionic signal of the peptides depended on the mass the analyte, while at higher pH values the current

signal of negatively charged peptides was higher than expected from their mass alone.

Without wishing to be bound by any theory, the inventors' explanation is that the peptides analyzed lost their charge, while the constriction of the nanopore still retained enough negative charge to recognize the peptide charge. Most likely, a negatively charged constriction is important for creating an electrophoretic environment for peptide-mass recognition. At the same time, the electrostatic interaction of the constriction with negatively charged analytes might prevent the correct position of the analyte within the reading frame of the nanopore.

Presumably, peptides need to be uniformly charged which can be achieved by lowering the pH of the solution. At the same time, however, the constriction of the nanopore should be negatively charged in order to obtain optimal mass recognition. Obtaining both effects may be challenging, because by lowering the pH the charge of the constriction also becomes less charged.

Therefore, in addition to the amino acid substitution(s) disclosed herein above, mutant FraC monomers may comprise one or more unnatural amino acids comprising a moiety that holds a negative charge at low pH, preferably wherein said moiety is a sulfate or phosphate group. In one embodiment, a reside (e.g. at position 10) is mutated to cysteine and then oxidized, hence introducing a sulfonic (or sulfinic) group at that position. The charge of such group remains negative over the all pH range. With this approach, the recognition of peptides could be improved. Alternatively, peptides might be chemically modified (e.g. by esterification) to neutralize the negative charge.

In a system according to the invention, the FraC nanopores are comprised in a lipid bilayer. The reconstitution of FraC nanopores in lipid bilayers has been described in the art. Typically, the lipid bilayer comprises phosphatidylcholine (PC), preferably 1,2-diphytanoyl-*sn*-glycero-3-

phosphocholine (DPhPC), optionally in combination with sphingomyelin (SM). Very good results are obtained when DPhPC and SM are present in about equal amounts by mass.

When a system according to the invention is in use, the nanopore is typically positioned between a first liquid medium and a second liquid medium, wherein at least one liquid medium comprises an analyte, and wherein the system is operative to detect a property of the analyte. In one embodiment, the system is operative to detect a property of the analyte comprises subjecting the nanopore to an electric field such that the analyte electrophoretically and/or electroosmotically translocates through the nanopore. As exemplified herein below, a system provided herein is particularly suitable for the analysis of a proteinaceous substance, preferably a peptide, more preferably a peptide up to about 30 amino acids in length. More in particular, a system of the invention provides for capture of peptides with different charge, recognition of the mass of the peptide and a resolution up to only 40 Da.

However, this is in no way to be understood that the invention is limited to applications relating to peptide analysis. For example, other analytes that can be detected using a system of the invention include (non-proteinaceous) biomarkers, antibiotics or other drugs, DNA, metabolites and small biological molecules.

The invention further relates to a mutant Fragaceatoxin C (FraC) polypeptide comprising one or more of the above mutations. These polypeptides are advantageously used in an (analytical) system herein disclosed. Also provided is an isolated nucleic acid molecule encoding a mutant FraC polypeptide according to the invention, and an expression vector comprising the isolated nucleic acid molecule. Still further, the invention provides a host cell comprising said expression vector.

In one embodiment, the mutant FraC polypeptide comprises a mutation at position D109, W112 and/or W116. For example, it comprises a mutation at W112 (optionally while W116 is not mutated) or it comprises a mutation at W116 (optionally while W112 is not mutated). In one embodiment, it comprises a mutation at both W112 and W116. As indicated herein above, the mutation(s) may comprise the substitution of D or W with S, T, A, N, Q or G, preferably with S or T. In a specific aspect, the invention provides mutant FraC W112S, FraC W116S, or FraC W112S/ W116S, or its encoding nucleic acid molecule, or vector comprising the same. Still further, it provides a polypeptide comprising mutation D109S, preferably wherein the mutant is D109S/W116S, or its encoding nucleic acid molecule, or vector comprising the same.

Any one of these mutations may be supplemented with one or more unnatural amino acids comprising a moiety that holds a negative charge at low pH, for example wherein said moiety is a sulfate, sulfonic acid or phosphate group. Preferred positions for introducing such negative charge residue(s) include one or more of positions 10, 17 and 24.

In a specific aspect, the mutation W116 is supplemented with mutation D10C. The thiol group of the cysteine is then oxidized to sulfonic acid e.g. by incubation of FraC double mutant monomers with 10% (v/v) hydrogen peroxide. It was found that the introduction of a sulfonic acid moiety at position 10 of FraC gives rise to oligomerised pores that show a quiet signal in electrophysiology recordings as compared to a more noisy signal observed for nanopores that had not been subjected to oxidation. Accordingly, in one embodiment the invention provides a mutant comprising the D10C substitution, preferably in combination with one or more of W112S, W116S and D109S, more preferably in combination with at least W116S.

A further embodiment relates to a method for providing a system according to the invention, comprising the steps of

- providing recombinant FraC monomers;
- contacting said monomers with liposomes to assemble them into oligomers;
- recovering the oligomers from the liposomes; and
- contacting the oligomers with a lipid bilayer, which may contain sphingomyelin, to allow the formation of FraC nanopores.

In one embodiment, the contacting with a lipid bilayer is performed at a pH below 4.5, preferably below 4.0.

A method of the invention may furthermore comprise the step of isolating a fraction comprising FraC nanopores in the Type II state, and/or a fraction comprising FraC nanopores in the Type III state. In one aspect, it comprises isolating different oligomeric forms of FraC using a liquid chromatographic technique, including size-exclusion, affinity, reverse-phase or ion exchange chromatography.

A peptide mass-detecting FraC nanopore system of the present invention is advantageously integrated in real-time protein sequencing system. To that end, the system preferably comprises one or more further modifications.

In one embodiment, a protease-unfoldase pair is attached directly above (i.e. on the *cis* side of) the FraC nanopore. Then, cleaved peptides will be sequentially recognized and translocated across the nanopore. For example, the barrel-shaped ATP-dependent ClpXP protease is an ideal candidate because it can encase the digested peptides preventing its release in solution. Another approach is based on a protein complex that constitutes the proteasome, or any other protease. For example, the complex includes the 20S alpha/beta subunits of the proteasome, and the 19S regulatory particle, of which the ATPase is the minimal required unit. However, other proteases could also be used. The protease will cleave the polypeptide specifically (for example it could cut after a positively charged residue or a negatively charged residue or a aromatic residue etcetera), or it will be engineered to cut specifically, or it will cut at nonspecific locations within

the polypeptide chain. The protease will ideally encase the substrate and will allow the docking of other components (e.g. unfoldases) to feed the polypeptide to the unfoldase active site.

The attachment may be of a covalent or non-covalent nature. For example, it can achieved by chemical attachment, by genetic fusion, or by introducing a binding loop into the FraC nanopore that can interact non-covalently with the peptidase.

We demonstrated that the peptides entering the *cis* side of the nanopore have a high probability of exiting the nanopore to the *trans* chamber, which will prevent duplicate detection events. Furthermore, we showed that at low pH peptides are likely to be captured and their mass recognized by the nanopore at a fixed applied potential irrespectively of their chemical composition. If such low pH values will not be compatible with enzymatic activity, asymmetric solutions on both side of the nanopore can be used[34,35]. In such system, conditions in the *cis* side can be tuned to optimize the ATPase activity of the unfoldase-peptidase, while the pH and ionic strength of the *trans* side can be optimized to capture and recognize individual peptides.

A FraC nanopore or a mutant FraC polypeptide as provided herein is advantageously used in all sorts of analyte analysis, including peptide or DNA analysis, preferably wherein peptide analysis comprises peptide mass detection and/or peptide sequencing. However, whereas the advantageous properties of a system provided herein are demonstrated in the context of peptide analysis, a person skilled in the art will appreciate that it can be used for various applications. Other possible applications of the invention include the following:

- Peptide post-modification detection (glycosylation etc), proteomics;
- DNA sequencing with higher accuracy, or DNA post-modification detection (methylation etc);

- Other small analytes detection or polymers analysis with higher resolution.
- Single molecule protein and DNA sequencing;
- Directly peptides, biomarkers, antibiotics and small molecules detection in human samples;
- Trapping of different size proteins for binding ligands analysis like glucose etc.

Mass spectrometry is the workhorse of the proteomics field. At present, the nanopore system falls short from the resolution of commercial mass spectrometers. A peptide mass-analyzer device based on FraC nanopores as herein disclosed has distinctive advantages compared to conventional mass spectrometers, which are expensive, extremely complex and unwieldy, and are not single-molecule. By contrast, nanopores can be integrated in portable and low-cost devices containing hundreds of thousands of individual sensors. Hence, in one embodiment the FraC nanopore system is integrated in a portable device comprising a plurality of individual FraC nanopore systems as herein described.

In addition, the electrical nature of the signal allows sampling biological samples in real-time. Furthermore, since the nanopore reads individual molecules, the signal contains additional information not available for ensemble techniques. In particular, single-molecule detection, especially when coupled to high throughput analysis, will allow detecting low abundance peptides and to unravel the chemical heterogeneity in post translational modifications, challenges that are hard to address with conventional mass spectrometry. For example, the invention also provides for the use of a FraC nanopore system or a mutant FraC polypeptide in single molecule detection, preferably in combination with high throughput analysis.

LEGEND TO THE FIGURES

Figure 1. Preparation and characterization of type I, type II and type III FraC nanopores. a, Cut through of a surface representation of WT-FraC oligomer (PDB: 4TSY)[25] colored according to the vacuum electrostatic potential as calculated by Pymol. One protomer is shown as a carton presentation with tryptophans 112 and 116 displayed as spheres. **b,** Percentage of the distribution of type I, type II and type III for WT-FraC, W112S-FraC, W116S-FraC and W112S-W116S-FraC at pH 7.5 and 4.5. **c,** IV curves of type II nanopores formed by WT-FraC, W116S-FraC and W112S-W116S-FraC at pH 7.5 (15 mM Tris.HCl, 1 M KCl). **d,** Single nanopore conductance of W116S-FraC in 1 M KCl (0.1 M citric acid and 180 mM Tris base) at pH 4.5. **e,** Typical current traces for the three nanopore types of W116S-FraC in 1 M KCl at pH 4.5 under -50 mV applied potential. **f,** Reversal potentials measured under asymmetric condition of KCl (1960 mM *cis*, 467 mM *trans*) at pH 4.5 for the three W116S-FraC nanopore types. The ion selectivity was given by using the Goldman–Hodgkin–Katz equation (equation 1). **g,** Molecular models of the three type FraC nanopores constructed from the FraC crystals structure using the symmetrical docking function of Rosetta. The electrophysiology recordings were performed with 10 kHz sampling and 2 kHz filter. The error bars and color shadow in the I-V curves are standard deviations from three repeats at least.

Figure 2. Single channel conductance distributions of FraC nanopores at pH 7.5 and 4.5. a, The table reports the average conductance values which were obtained by fitting Gaussian functions to conductance histograms. S.D. represents the standard deviation of all single channels (number given as n). **b-f**, Each panel represents a different batch of FraC nanopores as indicated. Single channels were collected under -50 mV applied potential. pH 7.5 and 4.5 were obtained using 1 M NaCl, 15 mM Tris, or 1 M KCl, 0.1 M citric acid, 180 mM Tris base respectively.

Figure 3. Discrimination of angiotensin peptides in mixture with type II W116S-FraC nanopores. a, (**i**) Sequences of angiotensin I, II, III and IV with corresponding Ires% measured at -30 mV. (**ii**) Typical blockades provoked by the four angiotensin peptides. (**iii**) Color density plot of the Ires% versus the standard deviation of the current amplitude for angiotensin I added to the *cis* compartment, and after the further addition of angiotensin II, angiotensin III, and angiotensin IV to the *cis* chamber (**iv**). **b,** Discrimination of angiotensin II and angiotensin A. (**i**) Table showing the sequences, the molecular weights and the Ires% of the peptides. The peptides differ by one amino acid as shown in red. (**ii**) Representative traces of the peptide blockades. Color density plot of the Ires% versus the standard deviation of the current amplitude for angiotensin II blockades prior (**iii**) and after (**iv**) the further addition of angiotensin A to the *cis* chamber. All measurements and recordings were performed in pH 4.5 buffer containing 1 M KCl, 0.1 M citric acid, 180 mM Tris base with a 50 kHz sampling and 10 kHz filter. Standard deviations were calculated from minimum three repeats. Color density plot was created with Origin.

Figure 4. Evaluation of biological peptides having different chemical compositions. Relation between the molecular weight and Ires% of peptide using: (**a**) type I WT-FraC nanopores, (**b**) type II W116S-FraC nanopores and (**c**) type III W112S-W116S-FraC nanopores at pH 4.5. The solid line represents a second order polynomial fitting. Current blockades were measured at -30 mV for type I and II pore, and at -50 mV for type III pore. Error bars are standard deviations obtained from at least three measurements.

Figure 5. A nanopore peptide mass spectrometer at pH 3.8. a, Amino acid sequences of four different peptides and their overall charge at different pH. The chargeable amino acids are underlined. **b,** pH dependence of the Ires% for the four peptides (*cis*) shown in **a** using type II W116S-FraC nanopores under -30 mV applied potential. **c**, Comparison of the Ires%

versus the mass of peptides at pH 4.5 and 3.8. **d**, Voltage dependence of c-Myc dwell times at different pHs. All electrophysiology measurements were carried out in 1 M KCl, 0.1 M citric acid, and pH was adjusted with 1 M Tris base to desired values. 50 kHz sampling rate and 10 kHz filter was used for collecting the current events. Error bars are standard deviations obtained from at least three measurements. The charges of the peptides were calculated according to the pKa for individual amino acids[36].

Figure 6: Discrimination of short peptide mixture with type III FraC nanopores comprising mutant W112S-W116S-FraC. a, Sequence, Ires% (-50 mV) and MW of angiotensin IV, angiotensin 4-8, endomorphin I and leucine enkephalin. **b**, Typical blockades provoked by the different peptides. **c**, Color density plot showing the Ires% versus the standard deviation of the current blockade for the mixture of angiotensin IV, angiotensin 4-8, endomorphin I and leucine-enkephalin. All measurements and recordings were performed in pH 4.5 buffer containing 1 M KCl, 0.1 M citric acid, 180 mM Tris base with a 50 kHz sampling and 10 kHz filter. Standard deviations were calculated from three repeats at least.

Figure 7. Characterization of type II FraC nanopores comprising an oxidized cysteine at position 10. Difference between the D10C/ W116S type II pore (panel A) and the oxidized D10C / W116S type II pore (panel B). Recordings were performed in a buffer containing 1 M NaCl, pH 7.5, ± 50 mV.

EXPERIMENTAL SECTION

Materials and Methods

Chemicals

Endothelin 1 (≥97%, CAS# 117399-94-7), endothelin 2 (≥97%, CAS# 123562-20-9), dynorphin A porcine (≥95%, CAS# 80448-90-4), angiotensin I (≥90%, CAS# 70937-97-2), angiotensin II (≥93%, CAS# 4474-91-3), c-Myc 410-419 (≥97%, # M2435), Asn1-Val5-Angiotensin II (≥97%, CAS# 20071-00-5), Ile7 Angiotensin III (≥95%, #A0911), leucine enkephalin (≥95%, #L9133), 5-methionine enkephalin (≥95%, CAS# 82362-17-2), endomorphin I (≥95%, CAS# 189388-22-5), pentane (≥99%, CAS# 109-66-0), hexadecane (99%, CAS# 544-76-3), Trizma®hydrochloride (≥99%, CAS# 1185-53-1), Trizma®base (≥99%, CAS# 77-86-1), Potassium chloride (≥99%, CAS# 7447-40-7), *N,N*-Dimethyldodecylamine *N*-oxide (LADO, ≥99%, CAS# 1643-20-5) were obtained from Sigma-Aldrich. Pre angiotensin 1-14 (≥97%, # 002-45), angiotensin 1-9 (≥95%, # 002-02), angiotensin A (≥95%, # 002-36), angiotensin III (≥95%, # 002-31), angiotensin IV (≥95%, # 002-28) were purchased from Pheonix Pharmaceuticals. Angiotensin 4-8 (≥95%) was synthesized by BIOMATIK. *1,2*-diphytanoyl-*sn*-glycero-*3*-phosphocholine (DPhPC, #850356P) and sphingomyelin (Porcine brain, # 860062) were purchased from Avanti Polar Lipids. Citric acid (99.6%, CAS# 77-92-9) was obtained from ACROS. *n*-Dodecyl β-D-maltoside (DDM, ≥99.5%, CAS# 69227-93-6) was bought from Glycon Biochemical EmbH. DNA primers were synthesized from Integrated DNA Technologies (IDT), enzymes from Thermo scientific. All peptides were dissolved with Milli-Q water without further purification and stored in -20°C freezer. pH 7.5 buffer containing 15 mM Tris in this study was prepared by dissolving 1.902 g Trizma® HCl and 0.354 g Trizma® base in 1 litre Milli-Q water (Millipore, Inc).

FraC monomer expression and purification

FraC gene containing *NcoI* and *HindIII* restriction site at the 5' and 3' ends, respectively, and a sequence encoding for a poly histidine tag at the 3' terminus was cloned to a pT7-SC1 plasmid. Plasmids were transformed into BL21(DE3) *E.cloni*® competent cell by electroporation. Cells were grown on LB agar plate containing 100 μg/mL ampicillin overnight at 37°C. The entire plate was then harvested and inoculated into 200 mL fresh 2YT media and the culture was grown with 220 rpm shaking at 37 °C until the optical density at 600 nm of the cell culture reached 0.8. Then 0.5 mM IPTG was added to the media and the culture was transferred to 25 °C for overnight growth with 220 rpm shaking. The next day the cells were centrifuged (2000 x g, 30 minutes) and the pellet stored at -80 °C. Cell pellets harvested from 100 mL culture media were used to purify FraC monomer. 30 mL of cell lysis buffer (150 mM NaCl, 15 mM Tris, 1 mM $MgCl_2$, 4 M urea, 0.2 mg/mL lysozyme and 0.05 unit/mL DNase) was added to resuspend the pellet and vigorously mixed for 1 hour. Cell lysate was then sonicated with Branson Sonifier 450 for 2 minutes (duty cycle 10%, output control 3). Then the crude lysate was centrifuged down at 4 °C for 30 minutes (5400 x g), and the supernatant incubated with 100 μL Ni-NTA beads (Qiagen) for 1 hour with gentle shaking. Beads were spun down and loaded to a Micro Bio-spin column (Bio-rad). 10 mL of SDEX buffer (150 mM NaCl, 15 mM Tris, pH 7.5) containing 20 mM imidazole was used to wash the beads, and proteins were eluded with 150 μL elution buffer (SDEX buffer, 300 mM imidazole). The concentration of protein was measured by measuring the absorption at 280 nm with Nano-drop 2000 (Thermo scientific) using the elution buffer as blank. To further confirm the purity of monomer, monomeric FraC was diluted to 0.5 mg/mL using the elution buffer and 9 μL of the diluted sample was mixted with 3 μL of 4x loading buffer (250 mM Tris HCl, pH 6.8. 8% SDS, 0.01% bromophenol blue and 40% glycerol) and then loaded to 12% SDS-PAGE gel. Gels were run for 30 min with 35 mA constant applied current, and stained with coomassie dye

(InstantBlue™, Expdedeon) for more than 1 hour before viewing using a gel imager (Gel Doc™, Bio-rad).

≥FraC | B9W5G6 (amino acid sequence)

SADVAGAVIDGAGLGFDVLKTVLEALGNVKRKIAVGIDNESGKTWTAM
NTYFRSGTSDIVLPHKVAHGKALLYNGQKNRGPVATGVVGVIAYSMSDG
NTLAVLFSVPYDYNWYSNWWNVRVYKGQKRADQRMYEELYYHRSPFR
GDNGWHSRGLGYGLKSRGFMNSSGHAILEIHVTKA

> 6xHis-WtFraC (amino acid sequence) as used in the present invention. Bold residues indicate residues of the N- and C-terminal end that were added to the original sequence.

MASADVAGAVIDGAGLGFDVLKTVLEALGNVKRKIAVGIDNESGKTWTA
MNTYFRSGTSDIVLPHKVAHGKALLYNGQKNRGPVATGVVGVIAYSMS
DGNTLAVLFSVPYDYNWYSNWWNVRVYKGQKRADQRMYEELYYHRSP
FRGDNGWHSRGLGYGLKSRGFMNSSGHAILEIHVTKA**GSAHHHHHH**

>6xHis-WtFraC (DNA sequence)

ATGGCGAGCGCCGATGTCGCGGGTGCGGTAATCGACGGTGCGGGTCTG
GGCTTTGACGTACTGAAAACCGTGCTGGAGGCCCTGGGCAACGTTAAA
CGCAAAATTGCGGTAGGGATTGATAACGAATCGGGCAAGACCTGGACA
GCGATGAATACCTATTTCCGTTCTGGTACGAGTGATATTGTGCTCCCAC
ATAAGGTGGCGCATGGTAAGGCGCTGCTGTATAACGGTCAAAAAAATC
GCGGTCCTGTCGCGACCGGCGTAGTGGGTGTGATTGCCTATAGTATGT
CTGATGGGAACACACTGGCGGTACTGTTCTCCGTGCCGTACGATTATAA
TTGGTATAGCAATTGGTGGAACGTGCGTGTCTACAAAGGCCAGAAGCG
TGCCGATCAGCGCATGTACGAGGAGCTGTACTATCATCGCTCGCCGTTT
CGCGGCGACAACGGTTGGCATTCCCGGGGCTTAGGTTATGGACTCAAA
AGTCGCGGCTTTATGAATAGTTCGGGCCACGCAATCCTGGAGATTCAC
GTTACCAAAGCAGGCTCTGCGCATCATCACCACCATCACTGATAAGCTT

FraC mutation preparation

FraC mutants were prepared according to MEGAWHOP method. 25 μL REDTaq® ReadyMix™ was mixed with 4 μM primer (see Table 1) containing the desired mutation with 50 ng plasmid (pT7-SC1 with wild type FraC gene) as template and the final volume was brought to 50 μL with MilliQ water.

Table 1. Primer sequences used in this study for preparing FraC mutants.

Primer name	DNA sequences
T7 promoter	5' TAATACGACTCACTATAGGG 3'
T7 terminator	5' GCTAGTTATTGCTCAGCGG 3'
W112S Fw	5' ACGATTATAATAGCTATAGCAATTGGTGG 3'
W116S Fw	5' ATTGGTATAGCAATAGCTGGAACGTG 3'
W112116S Fw	5' GTACGATTATAATAGCTATAGCAATAGCTGGAACGTGC 3'
D109S ReV	5' TGCTATACCAATTATAGCTGTACGGCA 3'

The PCR protocol was initiated by 150 seconds denature step at 95 °C, followed by 30 cycles of denaturing (95 °C, 15 s), annealing (55 °C, 15 s), and extension (72 °C, 60 s). The PCR products (MEGA primer) were combined and purified using a QIAquick PCR purification kit with final DNA concentration around 200 ng/μL. The second PCR was performed for whole plasmid amplification. 2 μL of MEGA primer, 1 μL Phire II enzyme, 10 μL 5x Phire buffer, 1 μL 10 mM dNTPs, were mixed with PCR water to 50 μL final volume. PCR started with pre-incubated at 98 °C (30 s) and then 25 cycles of denaturing (98 °C, 5 s), annealing (72 °C, 90 s), extension (72 °C, 150 s). When the PCR was completed, 1 μL Dpn I enzyme was added and the mixture kept at 37 °C for 1 hour. Then the temperature was raised to 65 °C for 1 minute to inactivate the enzyme. Products were then transformed into *E. cloni*® 10G cells (Lucigen) competent cell by electroporation. Cells were plated on LB agar plates containing 100 μg/mL

ampicillin and grew at 37 °C overnight. Single clones were enriched and sent for sequencing.

Sphingomyelin-DPhPC liposome preparation

20 mg sphingomyelin and 20 mg DPhPC (1,2-diphytanoyl-*sn*-glycero-3-phosphocholine) were dissolved in 4 mL pentane with 0.5% v/v ethanol and brought to a 50 mL round flask. The solvent was then evaporated by rotation and using a hair-dryer to warm-up the flask. After evaporation, the flask was kept at ambient temperature for an additional 30 minutes. The lipid film was resuspended with 4 mL SDEX buffer (150 mM NaCl, 15 mM Tris, pH 7.5) and the solution immersed in a sonication bath for 5 minutes. Liposome suspensions were stored at -20°C.

FraC oligomerization

FraC oligomerization was triggered by incubation of FraC monomers with sphingomelyin-DPhPC liposomes. Frozen liposome were thawed and sonicated in a water bath for one minute. FraC monomers were diluted to one mg/mL using SDEX buffer, and then 50 µL of FraC monomers were added to 50 µl of a 10 mg/mL liposome solution to obtain a mass ratio of 10:1 (liposome : protein). The lipoprotein solution was incubated at 37 °C for 30 min to allow oligomerization. Then 10 µl of 5% (w/v, 0.5% final) LADO was added to the lipoprotein solution to solubilize the liposomes. After clarification (typically 1 minute) the solution was transferred to a 50 mL Falcon tube. Then 10 mL of SDEX buffer containing 0.02% DDM and 100 µL of pre-washed Ni-NTA beads were added to the Falcon tube and mixed gently in shaker for 1 hour at room temperature. The beads were then spun down and loaded to a Micro Bio-spin column. 10 mL wash buffer (150 mM NaCl, 15 mM Tris, 20 mM imidazole, 0.02% DDM, pH 7.5) was used to wash the beads and oligomers eluded with 100 µL elution buffer (typically 200 mM EDTA, 75 mM NaCl, 7.5 mM Tris pH 7.5, 0.02% DDM). The FraC oligomers were stored at 4 °C. Under these conditions the nanopores are stable for several months.

W112S-W116S-FraC oligomer separation with His-Trap chromatography

200 µL of W112S-W116S-FraC monomers (3 mg/mL) were incubated with 300 µL of Sphingomyelin-DPhPC liposome (10 mg/mL) and kept at 4 °C for 48 hours after which 0.5% LADO (final concentration) was added to solubilize the lipoprotein. Then the buffer was exchanged to the 500 mM NaCl, 15 mM Tris, 0.01% DDM, 30 mM imidazole, pH 7.5 (binding buffer) using a PD SpinTrap G-25 column. W112S-W116S-FraC oligomers were then loaded to Histrap HP 1 mL column (General Electric) using an ÄKTA pure FPLC system (General Electric). The loaded oligomers were washed with 10 column volumes of 500 mM NaCl, 15 mM Tris, 0.01% DDM, 30 mM imidazole, pH 7.5, prior applying an imidazole gradient (from 30 mM to 1 M imidazole, 500 mM NaCl, 15 mM Tris, 0.01% DDM, pH 7.5) over 30 column volumes. The signal was monitored with the absorbance at 280 nm and fractions were collected when the absorbance was higher than 5 mAu.

Electrophysiology measurement and data analysis

Electrical recordings were performed as explained in details previously[27,37]. I_O, referring to open pore current, were measured by fitting Gaussian functions to event amplitude histograms. Residual current values (Ires%) were calculated by dividing the blockade current (I_B) by open pore current ($I_B/I_O \times 100\%$). Dwell times and inter-event times were measured by fitting single exponentials to histograms of cumulative distribution.

Ion permeability measurement

In order to measure reversal potentials, a single channel was obtained under symmetric conditions (840 mM KCl, 500 µL in each electrophysiology chamber) and the electrodes were balanced. The 400 µL of a buffered stock solution of 3.36 M KCl was slowly added to *cis* chamber, while 400 µL of salt free buffered solution was added to the *trans* chamber to obtain a total volume of 900 µL (*trans*:*cis*, 467 mM KCl:1960 mM KCl). After the equilibrium was reached, IV curves were collected from -30 to + 30 mV. The

resulting voltage at zero current is the reversal potential (V*r*). The ion selectivity (P_{K^+}/P_{Cl^-}) was then calculated using the Goldman-Hodgkin-Katz equation (equation 1) where $\left[a_{K^+/Cl^-}\right]_{cis/trans}$ is the activity of the K^+ or Cl^- in the *cis* or *trans* compartment, *R* the gas constant, *T* the temperature and *F* the Faraday's constant.

$$\frac{P_{K^+}}{P_{Cl^-}} = \frac{[a_{Cl^-}]_{trans} - [a_{Cl^-}]_{cis} e^{V_r F/RT}}{[a_{K^+}]_{trans} e^{V_r F/RT} - [a_{K^+}]_{cis}} \qquad (1)$$

The activity of ions was calculated by multiplying the molar concentration of the ion for the mean ion mobility (0.649 for 500 mM KCl, and 0.573 for 2000 mM). Ag/AgCl electrodes were surrounded by 2.5% agarose bridge in 2.5 M NaCl.

Molecular models of Type I, II and III FraC nanopores

The 3D models with different multimeric order, ranging from five to nine monomers, were constructed with the symmetrical docking function of Rosetta[38]. A monomer without lipids was extracted from the crystal structure of FraC with lipids (PDB_ID 4tsy). Symmetrical docking arranged this monomer around a central rotational axis ranging in order form 5 to 9. In total Rosetta generated and scored 10 000 copies for each symmetry. In all cases, a multimeric organization with a symmetry similar to the crystal structure could be identified as a top scoring solution. However, in the pentameric assembly the multimer interface was not fully satisfied as compared to the crystal structure, with large portions left exposed. The 9-fold symmetric model however exhibited a significant drop in Rosetta score compared to the 6- 7- and 8-fold symmetric models indicating an unfavored assembly of the nonameric assembly with the 6- 7- and 8-fold assemblies as the most plausible. To create lipid bound models, the crystal structure with lipids was superimposed on each monomer of the generated models, allowing the lipid coordinates to be transferred. The residues within 4.5 angstrom of the lipids were minimized with the Amber10 forcefield.

EXAMPLE 1: Engineering the size of FraC Nanopores

One of the main challenges in biological nanopores analysis is to obtain nanopores with different size and shape. Most of biological nanopores are formed by multiple repeats of individual monomers. Hence, different nanopore sizes might be obtained by engineering the protein oligomeric composition[28]. We noticed that at pH 7.5 a small fraction of Wild Type FraC (WT-FraC) nanopores showed a lower conductance (1.26 ± 0.08 nS, -50 mV, type II WT-FraC, Figure 1b) compared to the dominant fraction (2.26 ± 0.08 nS, -50 mV, type I WT-FraC), suggesting that FraC might be able spontaneously to assemble into nanopores with smaller size. At pH 4.5 yet a smaller nanopore conductance was observed (0.42 ± 0.03 nS, type III WT-FraC, -50 mV, Figure 1b). We noticed that the reconstitution of lower conductance nanopores depended to several purification conditions. In particular, we observed that the occurrence of type II and type III nanopores increased when the oligomers were stored in solution for several weeks or when the concentration of monomeric WT-FraC was reduced during oligomerisation.

In an attempt to enrich for type II and type III FraC nanopores, the interaction between the nanopore and the lipid interface was weakened by substituting W112 and W116 at the lipid interface of FraC (Figure 1a) with serine. The inventors reasoned that a lower concentration of monomers during liposome-triggered oligomerisation would increase the population of lower molecular mass oligomers. Surprsingly, it was found that at pH 4.5 using W116S-FraC and W112S-W116S-FraC oligomers, type II and type III FraC nanopores were the dominant species, respectively (Figure 1b, Figure 2). Conveniently, the different nanopore types could be separated by Ni-NTA affinity chromatography using an imidazole gradient. Furthermore, it was found that enrichment of type II and type III FraC nanopores could also be obtained at pH 7.5 by replacing aspartate 109 at the lipid interface with serine (see Figure 2e, Table 2).

Table 2: relative amounts of Type I, Type II and Type III for each of the FraC nanopores investigated at neutral and acidic pH.

		Type I(%)	SD	Type II(%)	SD	Type III(%)	SD
pH 7.5	*Wild type*	85.7	3.8	14.3	3.8	0.0	0.0
	W112S	61.9	4.3	38.1	4.3	0.0	0.0
	W116S	61.1	5.7	38.9	5.7	0.0	0.0
	W112116S	27.1	3.9	72.9	3.9	0.0	0.0
	D109S	50.3	3.8	48.0	3.6	1.7	1.5
	D109SW116S	29.3	9.5	66.7	8.4	4.0	4.0
pH 4.5	*Wild type*	42.5	10.6	51.9	7.7	5.6	7.9
	W116S	29.0	4.0	47.0	4.6	24.0	2.3
	W112SW116S	21.7	4.7	38.0	8.5	40.3	9.3
	D109S	35.7	2.1	56.3	9.1	8.0	7.0
	D109SW116S	19.3	8.4	64.3	6.7	16.3	3.2

Among FraC nanopores of the same type, the lipid interface modifications caused by W112S and W116S substitutions did not alter the conductance and ion selectivity as compared to that of wild type (Figure 1c, Figure 2, Table 3) suggesting that the modifications did not altered the overall fold of the nanopores. When characterised in lipid bilayers, type I, type II and type III nanopores showed a well-defined single conductance distribution and a steady open pore current (Figure 1d-e). Interestingly, Type I, Type II and Type III nanopores showed increasing cation selectivity (from 2.0 for type I to 4.2 for type III W116S-FraC nanopores at pH 4.5 (Figure 1f, Table 3), most likely reflecting a larger overlap of the electrical double layer in the nanopores with a narrower constriction.

Table 3: Ion selectivity of different FraC pores at pH 7.5 and 4.5.

		pH 7.5		pH 4.5	
		Reversal potential (mV)	P_{K^+} / P_{Cl^-}	Reversal potential (mV)	P_{K^+} / P_{Cl^-}
WT-FraC	Type I	17.2±1.2	3.6±0.4	10.5±1.4	2.1±0.2
	Type II	20.8±1.6	5.2±0.9	12.3±1.2	2.4±0.2
	Type III	/	/	20.6±1.1	5.0±0.6
W116S-FraC	Type I	/	/	10.1±0.9	2.0±0.1
	Type II	/	/	12.8±0.7	2.5±0.2
	Type III	/	/	18.8±0.5	4.2±0.2
W112S-W116S-FraC	Type I	/	/	8.8±1.2	1.9±0.2
	Type II	/	/	14.0±0.1	2.8±0.1
	Type III	/	/	20.1±0.6	4.8±0.3

The ion selectivity (P_{K^+}/P_{Cl^-}) was calculated from the reversal potential according to the Goldman-Hodgkin-Katz equation: $\frac{P_{K^+}}{P_{Cl^-}} = \frac{[a_{Cl^-}]_{trans} - [a_{Cl^-}]_{cis} e^{V_r F/RT}}{[a_{K^+}]_{trans} e^{V_r F/RT} - [a_{K^+}]_{cis}}$, where Vr is the reversal potential, P_{K^+}/P_{Cl^-} the ion selectivity, a the activity of ions and F the Farady constant. Electrophysiology recordings were carried out with 1960 mM KCl in the *cis* solution and 467 mM KCl in the *trans* solution. The activity of ions was calculated by multiplying the molar concentration of the ion for the mean ion mobility (0.649 for 500 mM KCl, and 0.573 for 2000 mM)[3]. Errors are given as standard deviations calculated from 3 experiments at least.

These findings strongly suggest that the three types of FraC nanopores represent nanopores with different protomeric compositions. Molecular modelling allowed predicting the diameter of type II (1.1 nm) and type III (0.8 nm) nanopores (Figure 1g); and revealed that type III FraC is most likely the biological nanopore with the smallest constriction known to date.

EXAMPLE 2: Identification of peptides containing single amino acid substitutions using type II or type III FraC nanopores as sensor

Type II FraC nanopores were used to sample a series of angiotensin peptides (which in blood regulate blood pressure and fluid balance. The peptides were added to the *cis* side of type II W116S-FraC nanopores and the induced ionic current blockades (I_B) was measured. Residual currents percent (Ires%, defined as I_B / I_O x 100) were used instead of current blockades because they provided more reliable values when comparing different nanopores. Results are shown in Figure 3a and Table 4.

Angiotensin I (DRVYIHPFHL, 1296.5 Da), showed the deepest blockade (Ires%= 8.8±0.2) and angiotensin IV (VYIHPF, 774.9 Da) the shallowest blockade (Ires%= 38.9±4.0). The residual current of angiotensin II (DRVYIHPF, 1046.2 Da, Ires%= 17.9±1.3) and angiotensin III (RVYIHPF, 931.1 Da, Ires%= 22.1±0.5) fell at intermediate values. When the four peptides were tested simultaneously, individual peptides could be readily discriminated (Figure 3a).

Table 4 : Peptide analysis using different types of FraC nanopores at pH 4.5 The electrophysiology solution contained 1 M KCl, 0.1 M citric acid, 180 mM Tris base at pH 4.5. Recordings were performed using a 50 kHz sampling and applying 10 kHz Bassel filter. Standard deviations were obtained for at least three measurements. The charges of the peptides were calculated according to the pK_a for individual amino acid[36].

Peptide	Sequence	Molecular weight (g/mol)	Charge pH 7.5	Charge pH 4.5	Ires% (I_B/I_O)% pH4.5	Dwell time (ms)
WT-FraC type I pore, -30 mV						
Endothelin 2	CSCSSWLDKECVYFCHLDIIW	2546.9	-2.15	0.36	6.1±1.8	104.0±29.9
Endothelin 1	CSCSSLMDKECVYFCHLDIIW	2491.9	-2.15	0.36	7.5± 0.5	19.73±1.95
Dynorphin A	YGGFLRRIRPKLKWDNQ	2147.5	3.76	4.48	15.1±2.6	3.68±0.76
Pre angiotensinogen	DRVYIHPFHLVIHN	1758.9	0.03	3.45	24.6±2.3	0.29±0.04
Angiotensin I	DRVYIHPFHL	1296.5	-0.06	2.46	43.4±0.9	0.15±0.04
W116S-FraC type II pore, -30 mV						

Angiotensin I	DRVYIHPFHL	1296.5	-0.06	2.46	8.8±0.2	0.54±0.01
c-Myc 410-419	EQKLISEEDL	1203.3	-3.24	-1.19	30.0±3.4	0.12±0.01
Angiotensin 1-9	DRVYIHPFH	1183.3	-0.06	2.46	14.0±0.2	0.37±0.04
Angiotensin II	DRVYIHPF	1046.2	-0.15	1.47	17.9±1.3	0.37±0.04
Asn1Val5 Angiotensin II	NRVYVHPF	1031.2	0.85	2.03	19.6±0.2	0.34±0.06
Angiotensin A	ARVYIHPF	1002.2	0.85	2.03	21.0±0.6	0.34±0.02
Angiotensin III	RVYIHPF	931.1	0.85	2.03	22.1±0.5	0.35±0.04
Ile7 Angiotensin III	RVYIHPI	897.1	0.85	2.03	24.3±0.4	0.19±0.05

Angiotensin IV	VYIHPF	774.9	-0.15	1.02	38.9±4.0	0.15±0.06
W112S-W116S-FraC type III pore, -50 mV						
Angiotensin IV	VYIHPF	774.9	-0.15	1.02	1.1±0.8	0.61±0.07
Angiotensin 4-8	YIHPF	675.8	-0.15	1.02	8.2±0.4	0.40±0.04
Endomorphin I	YPWF	610.7	-0.24	0.04	19.2±0.5	0.32±0.04
Met5 Enkephalin	YGGFM	573.7	-0.24	0.04	33.5±0.7	0.16±0.02
Leucine Enkephalin	YGGFL	555.6	-0.24	0.04	34.5±2.4	0.20±0.05

The resolution limit of the nanopore sensor was challenged by sampling a mixture of peptides. Remarkably, angiotensin II and angiotensin A, having an identical composition except for the initial amino acid (aspartate in angiotensin II vs. alanine in angiotensin A), appeared as distinctive peaks in Ires% plots (Figure 3b). Smaller differences in peptide mass, *e.g.* the 34 Da difference between phenylalanine and leucine in angiotensin III and Ile7 angiotensin III, were observed but not easily detected (data not shown), indicating the resolution of our system at ~40 Da. Smaller peptides such as angiotensin II 4-8 (YIHPF, 675.8 Da), endomorphin I (YPWF, 610.7 Da) or leucine enkephalin (YGGFL, 555.6 Da) translocated too quickly across type II W116S-FraC nanopores to be sampled. However, they could be readily measured using type III W112S-W116S-FraC nanopores (Table 4; Figure 6).

EXAMPLE 3: A nanopore mass spectrometer for peptides

Although the ability of biological nanopores to distinguish between known analytes is useful, a more powerful application would be the identification of peptide masses directly from ionic current blockades without holding prior knowledge of the analyte identity. In an effort to assess FraC nanopores as peptide mass analyzer, additional peptides were tested at pH 4.5 and 1 M KCl using type I, type II and type III FraC nanopores (Figure 4a-c, Table 4). Crucially, analytes with largely different charge compositions were included.

It was found that for most of peptides there was a direct correlation between the size and the residual current (Figure 4a-c). A notable exception was c-Myc 410-419 (1203.3 Da), an intentionally selected peptide because it includes a long stretch of negatively charged residues (Figure 5a). The overall negative charge of the peptide at pH 4.5 (see Table 4) was expected to have an effect on both peptide capture and recognition. c-Myc 410-419 could be readily captured at negative applied potentials (*trans*), indicating that the *cis* to *trans* electroosmotic flow across the

nanopore can overcome the electrostatic energy barrier opposing peptide capture. However, the Ires% of c-Myc 410-419 (30.0 ± 3.4) was significantly higher than the expected value (Figure 4b).

We reasoned that such anomaly might be due to the interaction between the acidic amino acids of the peptide and the negatively charged constriction of FraC nanopores. Thus, we lowered the pH solution to values where the aspartate and glutamate side chains in the peptides are expected to be protonated, hence become neutral (Figure 5a). Rewardingly, at pH 3.8, the signal corresponding to c-Myc 410-419 (1203.3 Da) fell between the signal of angiotensin I (1296.5 Da), and angiotensin II (1046.2 Da, Figure 5b, c). This indicates that, after losing its negative charges, the peptide blockades scaled with the expected mass of the peptides.

It has been assumed[1,30,31] and experimentally proven[32] that the voltage dependence of the average dwell time (τ_{off}) can report on the translocation of a molecule across a nanopore. Under a negative bias (*trans*) for positively charged peptides (added in *cis*) both electrophoretic and electroosmotic forces (from *cis* to *trans*) promote the entry and translocation[27] across the nanopore. For negatively charged peptides, such as c-Myc 410-419 at pH 4.5 (Figure 5a), the electroosmotic driving force must be stronger than the opposing electrophoretic force. The voltage dependence of τ_{off} was examined for c-Myc 410-419 at different pH values (Figure 5d). At pH 4.5 the peptide exhibited a maximum in τ_{off} at -50 mV, suggesting that at low potentials c-Myc 410-419 returns to the *cis* chamber (<50 mV), and at higher potentials (>50 mV) c-Myc 410-419 exits to the *trans* chamber. At pH 3.8 and lower, we observed a decrease in τ_{off} at higher potentials, indicating that c-Myc 410-419 crosses the membrane to the *trans* chamber.

As shown in Fig. 6, type III FraC nanopore can detect differences in peptide length down to 4 amino acids (mass around 500 Dalton) in a peptide mixture. It was also found that the residual current signal correlated well with the mass of peptides, suggesting that Type III can be used as a detector for a peptide having a mass down to ~500 Da .

EXAMPLE 4: Peptide mass identifier at pH 3.

This example shows that mutation D10C can be used as additional mutation to obtain a FraC pore showing a quiet signal in electrophysiology recordings.

Using mutant W116S as exemplary mutant, the aspartic acid at position 10 of FraC was converted to cysteine by site-directed mutagenesis. The thiol group of cysteine was then oxidized to sulfonic acid by incubation of FraC monomers with 10% hydrogen peroxide (v/v), which was dissolved in regular buffer (e.g. 10 mM Tris buffer pH 7.5, 150 mM NaCl). As a control, the double mutant was left without oxidation.

D10C/W116S FraC was oligomerized, and the oligomers tested in electrical recordings. Figure 7 shows the trace comparison between the D10C/ W116S pore and oxidized D10C/ W116S pore to demonstrate the difference after oxidization. Oligomerised pores from oxidized D10C / W116S FraC monomers showed a quiet signal in electrophysiology recordings, as compared to a more noisy signal observed for nanopores that had not been subjected to oxidation.

REFERENCES

1. Clarke, J. *et al.* Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* **4,** 265–270 (2009).

2. Derrington, I. M. *et al.* Nanopore DNA sequencing with MspA. *Proc. Natl. Acad. Sci.* **107,** 16060–16065 (2010).

3. Bezrukov, S. M., Vodyanoy, I., Brutyan, R. A. & Kasianowicz, J. J. Dynamics and free energy of polymers partitioning into a nanoscale pore. *Macromolecules* **29,** 8517–8522 (1996).

4. Robertson, J. W. F. *et al.* Single-molecule mass spectrometry in solution using a solitary nanopore. *Proc. Natl. Acad. Sci.* **104,** 8207–8211 (2007).

5. Baaken, G. *et al.* High-Resolution Size-Discrimination of Single Nonionic Synthetic Polymers with a Highly Charged Biological Nanopore. *ACS Nano* **9,** 6443–6449 (2015).

6. Aksoyoglu, M. A. *et al.* Size-dependent forced PEG partitioning into channels: VDAC, OmpC, and α-hemolysin. *Proc. Natl. Acad. Sci.* **113,** 9003–9008 (2016).

7. Krasilnikov, O. V., Rodrigues, C. G. & Bezrukov, S. M. Single polymer molecules in a protein nanopore in the limit of a strong polymer-pore attraction. *Phys. Rev. Lett.* **97,** 1–4 (2006).

8. Oukhaled, A. G., Biance, A. L., Pelta, J., Auvray, L. & Bacri, L. Transport of long neutral polymers in the semidilute regime through a protein nanopore. *Phys. Rev. Lett.* **108,** 1–4 (2012).

9. Zhao, Q., Jayawardhana, D. A., Wang, D. & Guan, X. Study of peptide transport through engineered protein channels. *J. Phys. Chem. B* **113,** 3572–3578 (2009).

10. Bacri, L. *et al.* Discrimination of neutral oligosaccharides through a nanopore. *Biochem. Biophys. Res. Commun.* **412,** 561–564 (2011).

11. Movileanu, L., Schmittschmitt, J. P., Scholtz, J. M. & Bayley, H. Interactions of peptides with a protein pore. *Biophys. J.* **89,** 1030–1045 (2005).

12. Mohammad, M. M., Prakash, S., Matouschek, A. & Movileanu, L. Controlling a single protein in a nanopore through electrostatic traps. *J. Am. Chem. Soc.* **130,** 4081–4088 (2008).

13. Piguet, F. *et al.* Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore. *Nat. Commun.* **9,** (2018).

14. Lamichhane, U. *et al.* Peptide translocation through the mesoscopic channel: Binding kinetics at the single molecule level. *Eur. Biophys. J.* **42,** 363–369 (2013).

15. Stefureac, R., Long, Y. T., Kraatz, H. B., Howard, P. & Lee, J. S. Transport of α-helical peptides through α-hemolysin and aerolysin pores. *Biochemistry* **45,** 9172–9179 (2006).

16. Chavis, A. E. *et al.* Single Molecule Nanopore Spectrometry for Peptide Detection. *ACS Sensors* **2,** 1319–1328 (2017).

17. Maglia, G., Restrepo, M. R., Mikhailova, E. & Bayley, H. Enhanced translocation of single DNA molecules through -hemolysin nanopores by manipulation of internal charge. *Proc. Natl. Acad. Sci.* **105,** 19720–19725 (2008).

18. Stoddart, D., Heron, A. J., Mikhailova, E., Maglia, G. & Bayley, H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc. Natl. Acad. Sci.* **106,** 7702–7707 (2009).

19. Boersma, A. J. & Bayley, H. Continuous stochastic detection of amino acid enantiomers with a protein nanopore. *Angew. Chemie - Int. Ed.* **51,** 9606–9609 (2012).

20. Stoddart, D. *et al.* Nucleobase recognition in ssDNA at the central constriction of the hemolysin pore. *Nano Lett.* **10,** 3633–3637 (2010).

21. Kennedy, E., Dong, Z., Tennant, C. & Timp, G. Reading the primary structure of a protein with 0.07 nm 3 resolution using a subnanometre-diameter pore. *Nat. Nanotechnol.* **11,** 968–976 (2016).

22. Li, S., Cao, C., Yang, J. & Long, Y. Detection of Peptides with Different Charges and Lengths by Using the Aerolysin Nanopore. **4,** 1–5 (2018).

23. Asandei, A. *et al.* Electroosmotic Trap Against the Electrophoretic Force Near a Protein Nanopore Reveals Peptide Dynamics during Capture and Translocation. *ACS Appl. Mater. Interfaces* **8,** 13166–13179 (2016).

24. Chinappi, M. & Cecconi, F. Protein sequencing via nanopore based devices: a nanofluidics perspective. *J. Phys. Condens. Matter* **in press,** (2018).

25. Tanaka, K., Caaveiro, J. M. M., Morante, K., González-Manãs, J. M. & Tsumoto, K. Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. *Nat. Commun.* **6,** 4–6 (2015).

26. Wloka, C., Mutter, N. L., Soskine, M. & Maglia, G. Alpha-Helical Fragaceatoxin C Nanopore Engineered for Double-Stranded and Single-Stranded Nucleic Acid Analysis. *Angew. Chemie - Int. Ed.* **55,** 12494–12498 (2016).

27. Huang, G., Willems, K., Soskine, M., Wloka, C. & Maglia, G. Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores. *Nat. Commun.* **8,** 1–13 (2017).

28. Soskine, M., Biesemans, A., De Maeyer, M. & Maglia, G. Tuning the size and properties of ClyA nanopores assisted by directed evolution. *J. Am. Chem. Soc.* **135,** 13456–13463 (2013).

29. Aqvist, J. *et al.* Dipoles Localized at Helix Termini of Proteins Stabilize Charges. *Proc. Natl. Acad. Sci.* **88,** 2026–2030 (1991).

30. Rincon-Restrepo, M., Mikhailova, E., Bayley, H. & Maglia, G. Controlled translocation of individual DNA molecules through protein nanopores with engineered molecular brakes. *Nano Lett.* **11,** 746–750 (2011).

31. Wanunu, M., Sutin, J., McNally, B., Chow, A. & Meller, A. DNA translocation governed by interactions with solid-state nanopores. *Biophys. J.* **95,** 4716–4725 (2008).

32. Biesemans, A., Soskine, M. & Maglia, G. A Protein Rotaxane Controls the Translocation of Proteins Across a ClyA Nanopore. *Nano Lett.* **15,** 6076–6081 (2015).

33. Ho, C. W. *et al.* Engineering a nanopore with co-chaperonin function. *Sci. Adv.* **1,** 1–9 (2015).

34. Wanunu, M., Morrison, W., Rabin, Y., Grosberg, A. Y. & Meller, A. Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. *Nat. Nanotechnol.* **5,** 160–165 (2010).

35. Stoddart, D., Franceschini, L., Heron, A., Bayley, H. & Maglia, G. DNA stretching and optimization of nucleobase recognition in enzymatic nanopore sequencing. *Nanotechnology* **26,** 10–16 (2015).

36. Stryer, L. *Biochemistry. Biochemistry (4th ed.)* (1995).

37. Soskine, M., Biesemans, A. & Maglia, G. Single-molecule analyte recognition with ClyA nanopores equipped with internal protein adaptors. *J. Am. Chem. Soc.* **137,** 5793–5797 (2015).

38. Andre, I., Bradley, P., Wang, C. & Baker, D. Prediction of the structure of symmetrical protein assemblies. *Proc. Natl. Acad. Sci.* **104,** 17656–17661 (2007).

Claims

1. A system comprising oligomeric Fragaceatoxin C (FraC) nanopores comprised in a lipid bilayer, wherein the sum of the nanopore fraction in the heptameric (Type II) state and the nanopore fraction in the hexameric (Type III) state represents at least 60% of the total number of FraC nanopores.

2. System according to claim 1, wherein the sum of the Type II and Type III state nanopores represents at least 65%, preferably at least 70%, of the total number of FraC nanopores.

3. System according to claim 1 or 2, wherein at least 60%, preferably at least 70%, of the FraC nanopores is in the Type II state.

4. System according to claim 1 or 2, wherein at least 60%, preferably at least 70%, of the FraC nanopores is in the Type III state.

5. System according to any one of the preceding claims, wherein the FraC nanopores comprise mutant FraC monomers comprising a mutation at position W112 and/or W116.

6. System according to claim 5, wherein said mutation(s) comprise substitution of W with S, T, A, N, Q or G, preferably with S or T.

7. System according to any one of the preceding claims, wherein the FraC nanopores comprise mutant FraC monomers comprising a mutation at position D109, preferably herein said mutation comprises the substitution of D with S or T, more preferably with S.

8. System according to any one of claims 5-7, wherein said mutant FraC monomer comprises mutation W112S, W116S and/or D109S, preferably wherein the FraC mutant is W112S/W116S or D109S/W116S.

9. System according to any one of the preceding claims, wherein the system is operative to detect a property of the analyte comprises subjecting the nanopore to an electric field such that the analyte electrophoretically and/or electroosmotically translocates through the nanopore.

10. System according to claim 9, wherein the analyte is a proteinaceous substance, preferably a peptide, more preferably a peptide up to about 30 amino acids in length.

11. A method for providing a system according to any one of claims 1 to 10, comprising the steps of:

(i) providing recombinant FraC monomers;

(ii) contacting said monomers with liposomes to assemble them into oligomers;

(iii) recovering the oligomers from the liposomes; (iv) contacting the oligomers with a lipid bilayer to allow the formation of FraC nanopores; and

(v) optionally isolating a fraction comprising FraC nanopores in the Type II state, and/or a fraction comprising FraC nanopores in the Type III state.

12. A mutant Fragaceatoxin C (FraC) polypeptide comprising a mutation at position W112 and/or W116 wherein the numbering corresponds to the FraC amino acid available under accession number B9W5G6 in UniProt, and wherein the W residue(s) is/are independently substituted with either S, T, A, N, Q or G.

13. A mutant Fragaceatoxin C (FraC) polypeptide comprising a mutation at position D109 wherein the numbering corresponds to the FraC amino acid available under accession number B9W5G6 in UniProt, and wherein the D residue is substituted with an uncharged residue, preferably with S or T, more preferably with S.

14. Mutant FraC polypeptide according to claim 12 or 13 comprising mutation W112S, W116S and/or D109S, preferably wherein the mutant FraC is W112S/W116S or D109S/W116S.

15. Mutant FraC polypeptide according to any one of claims 12 to 14, further comprising mutation D10C.

16. An isolated nucleic acid molecule encoding a mutant FraC polypeptide according to any one of claims 12 to 15.

17. An expression vector comprising an isolated nucleic acid molecule according to claim 16.

18. A host cell comprising an expression vector according to claim 17.

19. The use of a system according to any one of claims 1 to 10, or a mutant FraC polypeptide according to any one of claims 12 to 15, in peptide analysis, preferably wherein peptide analysis comprises peptide mass detection and/or peptide sequencing.

20. The use of a system according to any one of claims 1 to 10, or a mutant FraC polypeptide according to any one of claims 12 to 15, in single molecule detection, preferably in combination with high throughput analysis.

21. The use according to claim 19 or 20, wherein the system is integrated in a portable device comprising a plurality of individual systems according to any one of claims 1-10.

Figure 1

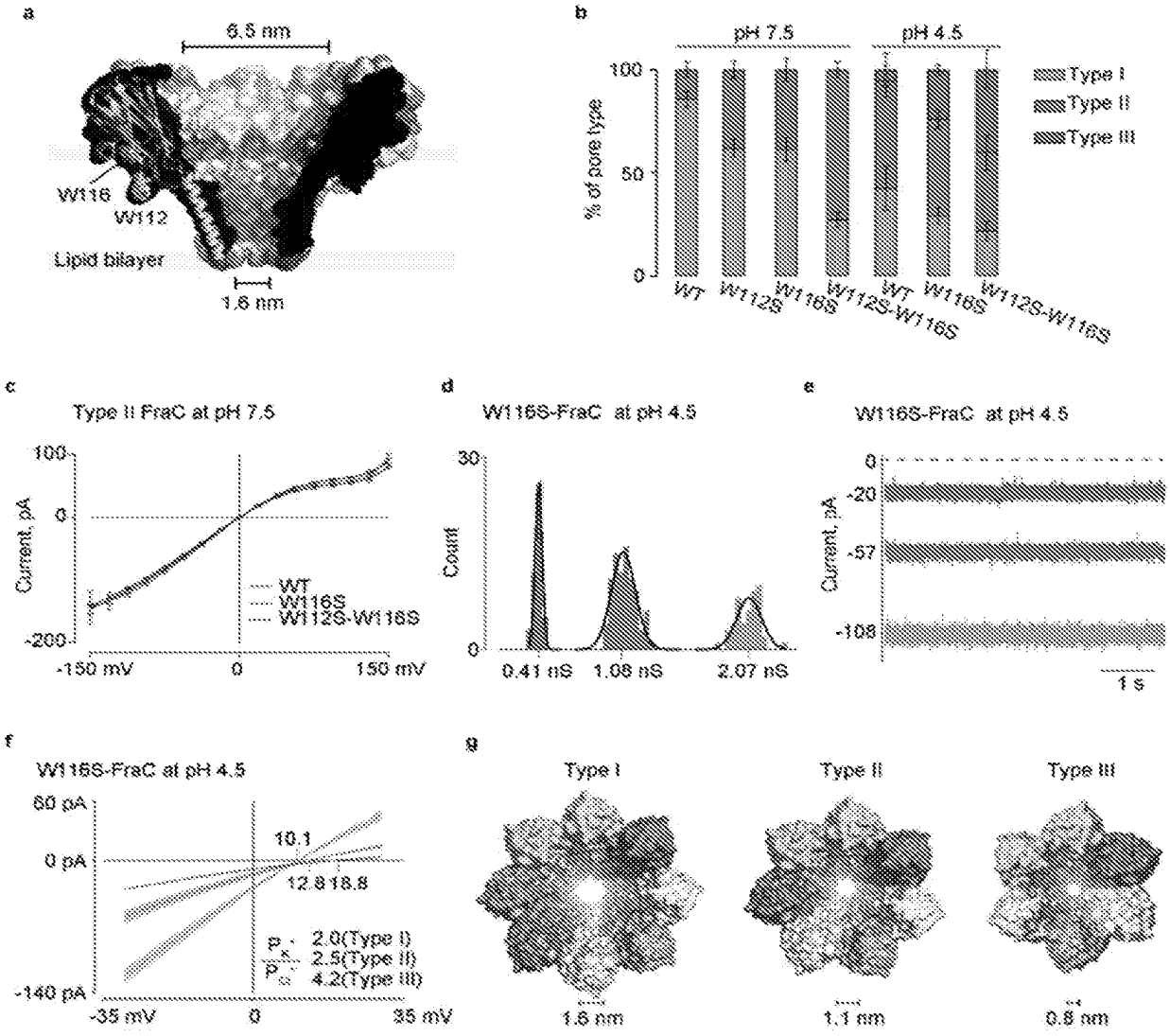
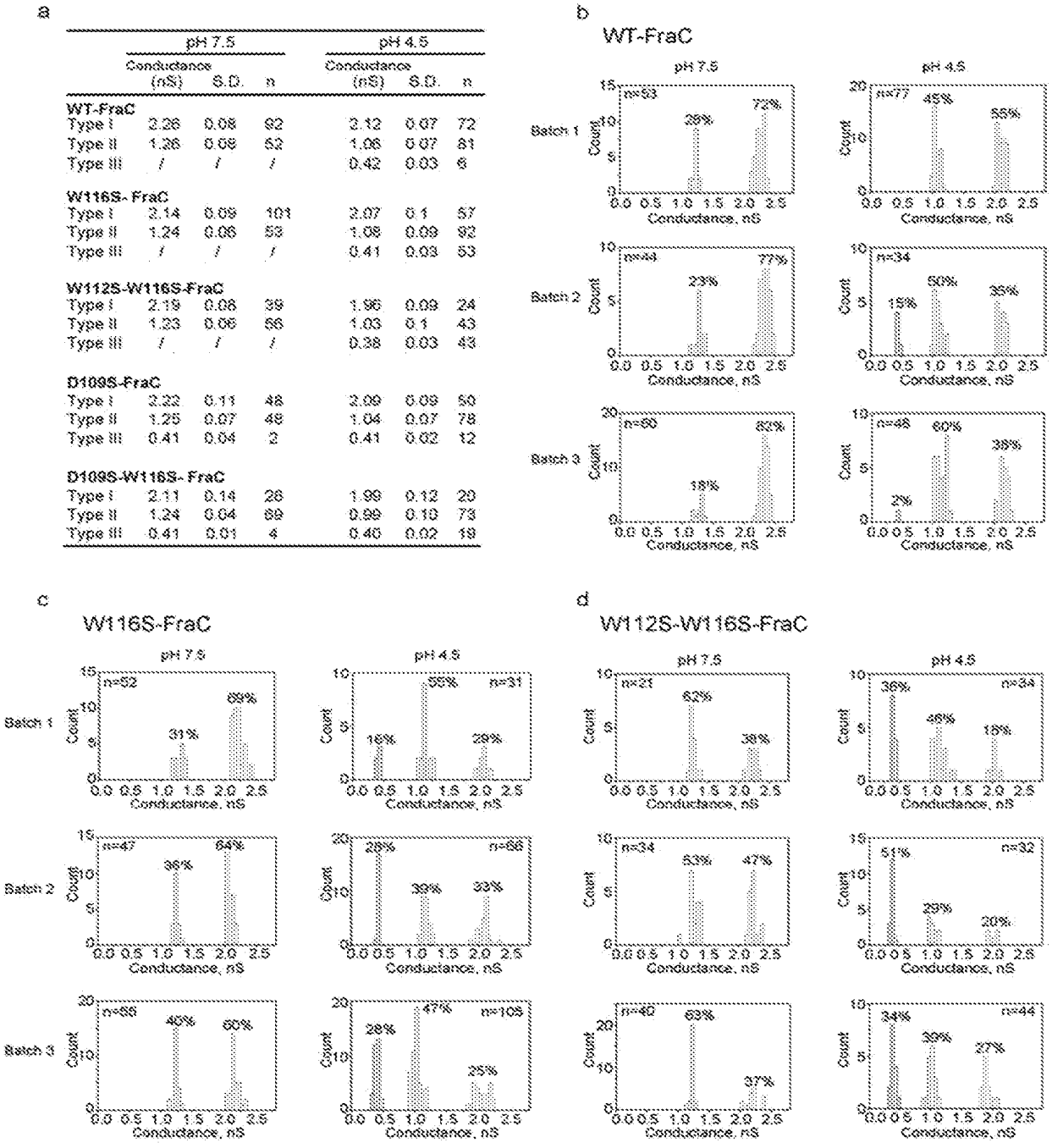

a
6.5 nm
W116
W112
Lipid bilayer
1.6 nm
b
pH 7.5
pH 4.5
100
50
0
% of pore type
Type I
Type II
Type III
WT
W112S
W116S
W112S-W116S
WT
W116S
W112S-W116S
c
Type II FraC at pH 7.5
100
0
-200
Current, pA
-150 mV
0
150 mV
WT
W116S
W112S-W116S
d
W116S-FraC at pH 4.5
30
0
Count
0.41 nS
1.08 nS
2.07 nS
e
W116S-FraC at pH 4.5
0
-20
-57
-108
Current, pA
1 s
f
W116S-FraC at pH 4.5
60 pA
0 pA
-140 pA
10.1
12.8 18.8
2.0(Type I)
2.5(Type II)
4.2(Type III)
-35 mV
0
35 mV
g
Type I
Type II
Type III
1.6 nm
1.1 nm
0.8 nm

Figure 2

a

	pH 7.5			pH 4.5		
	Conductance (nS)	S.D.	n	Conductance (nS)	S.D.	n
WT-FraC						
Type I	2.26	0.08	92	2.12	0.07	72
Type II	1.36	0.08	52	1.06	0.07	81
Type III	/	/	/	0.42	0.03	6
W116S- FraC						
Type I	2.14	0.09	101	2.07	0.1	57
Type II	1.24	0.06	53	1.08	0.09	92
Type III	/	/	/	0.41	0.03	53
W112S-W116S-FraC						
Type I	2.19	0.08	39	1.96	0.09	24
Type II	1.23	0.06	56	1.03	0.1	43
Type III	/	/	/	0.38	0.03	43
D109S-FraC						
Type I	2.22	0.11	48	2.09	0.09	50
Type II	1.25	0.07	48	1.04	0.07	78
Type III	0.41	0.04	2	0.41	0.02	12
D109S-W116S- FraC						
Type I	2.11	0.14	26	1.99	0.12	20
Type II	1.24	0.04	89	0.99	0.10	73
Type III	0.41	0.01	4	0.40	0.02	19



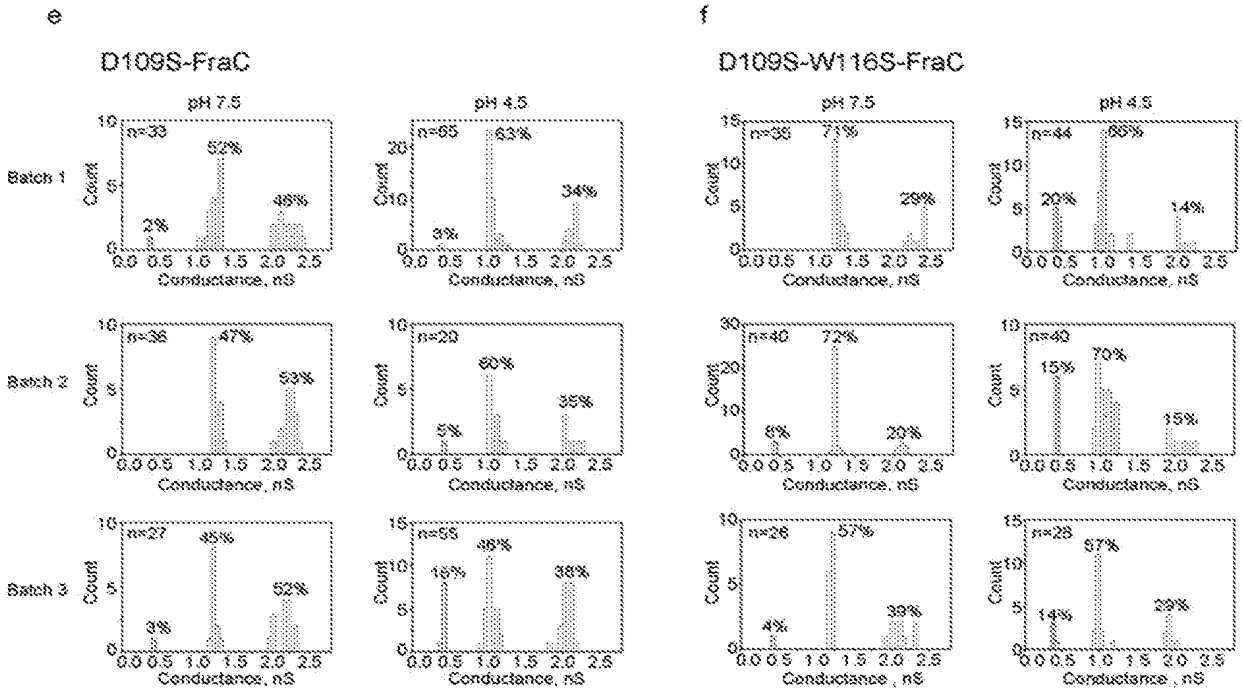

e
D109S-FraC
pH 7.5
pH 4.5
Batch 1
Batch 2
Batch 3
Count
Conductance, nS
n=33
52%
2%
46%
n=65
63%
3%
34%
n=36
47%
53%
n=20
60%
5%
35%
n=27
45%
3%
52%
n=58
46%
16%
38%
f
D109S-W116S-FraC
pH 7.5
pH 4.5
n=35
71%
29%
n=44
86%
20%
14%
n=40
72%
8%
20%
n=40
70%
15%
15%
n=26
57%
4%
39%
n=28
57%
14%
29%

Figure 3

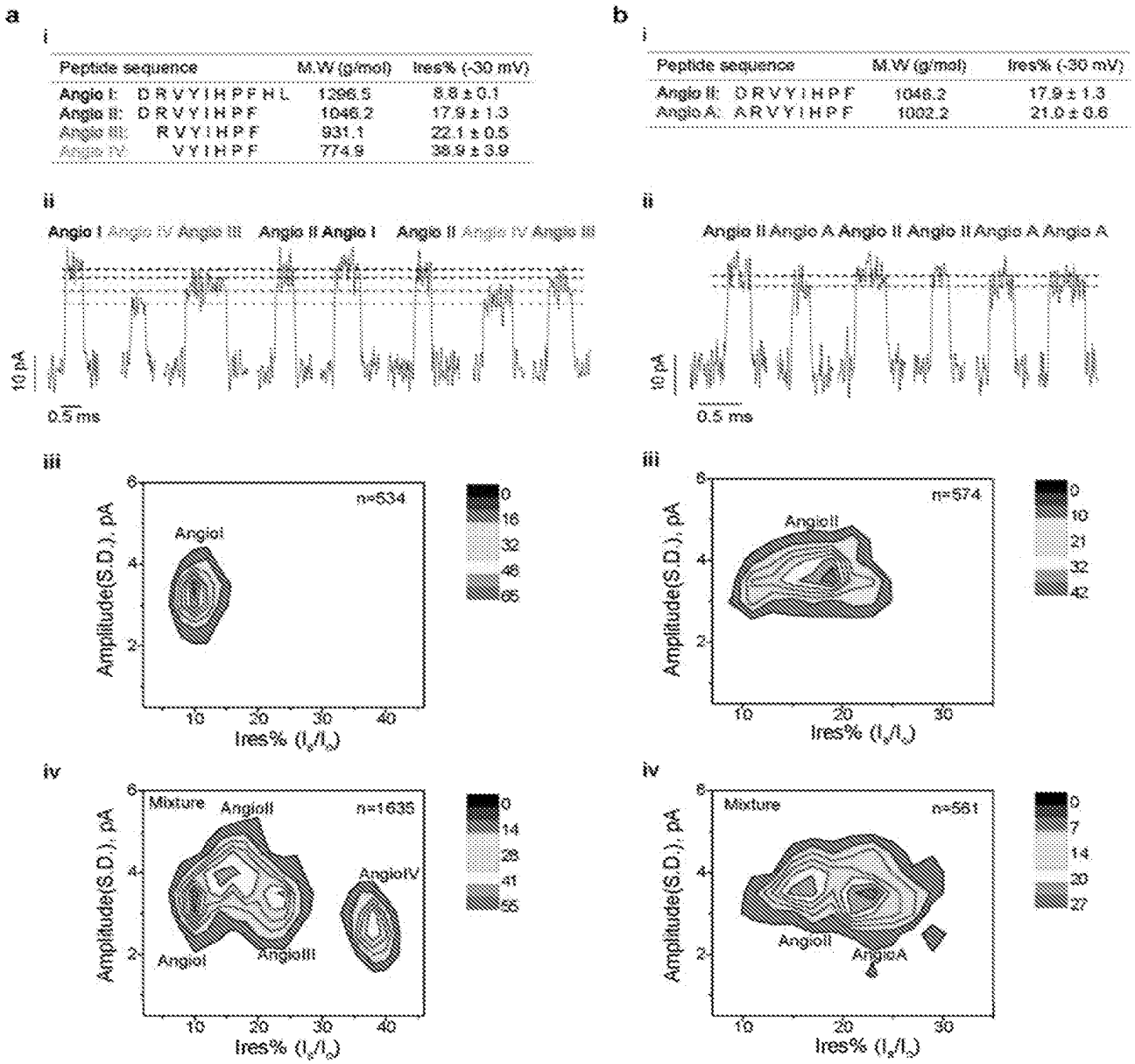

a
i
Peptide sequence M.W (g/mol) Ires% (-30 mV)
Angio I: D R V Y I H P F H L 1296.5 8.8 ± 0.1
Angio II: D R V Y I H P F 1046.2 17.9 ± 1.3
Angio III: R V Y I H P F 931.1 22.1 ± 0.5
Angio IV: V Y I H P F 774.9 38.9 ± 3.9
ii
Angio I Angio IV Angio III Angio II Angio I Angio II Angio IV Angio III
10 pA
0.5 ms
iii
n=634
Angiol
Amplitude(S.D.), pA
Ires% (Ix/I0)
iv
Mixture
AngioII
n=1635
AngioIV
AngioI
AngioIII
b
i
Peptide sequence M.W (g/mol) Ires% (-30 mV)
Angio II: D R V Y I H P F 1046.2 17.9 ± 1.3
Angio A: A R V Y I H P F 1002.2 21.0 ± 0.6
ii
Angio II Angio A Angio II Angio II Angio A Angio A
10 pA
0.5 ms
iii
n=574
AngioII
Amplitude(S.D.), pA
Ires% (Ix/I0)
iv
Mixture
n=561
AngioII
AngioA
Amplitude(S.D.), pA
Ires% (Ix/I0)

Figure 4

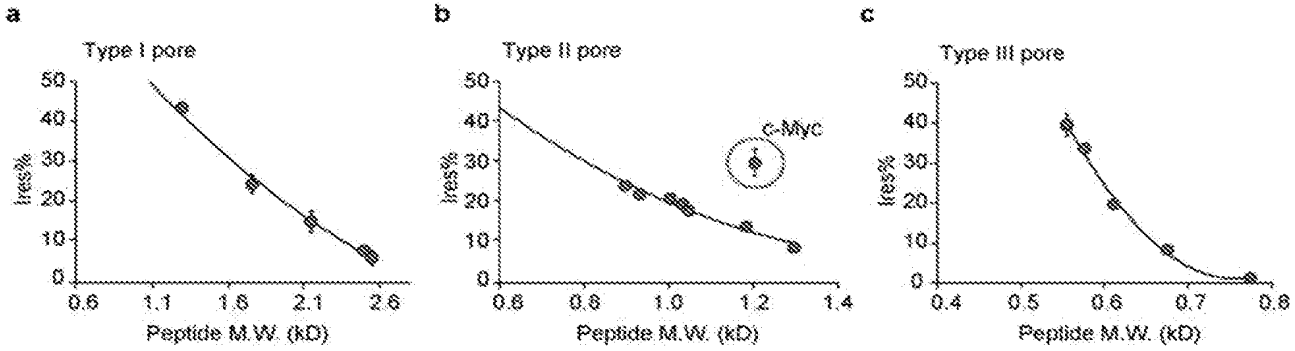


Figure 5

a

Peptide	Sequence	Calculated Charge pH 4.5	pH 3.8	pH 3.0
Ang I :	DRVYIHPFHL	+2.5	+3.0	+3.5
c-Myc :	EQKLISEEDL	-1.2	+0.4	+1.4
Ang II :	DRVYIHPF	+1.5	+2.0	+2.5
Ang III :	RVYIHPF	+2.0	+2.2	+2.6

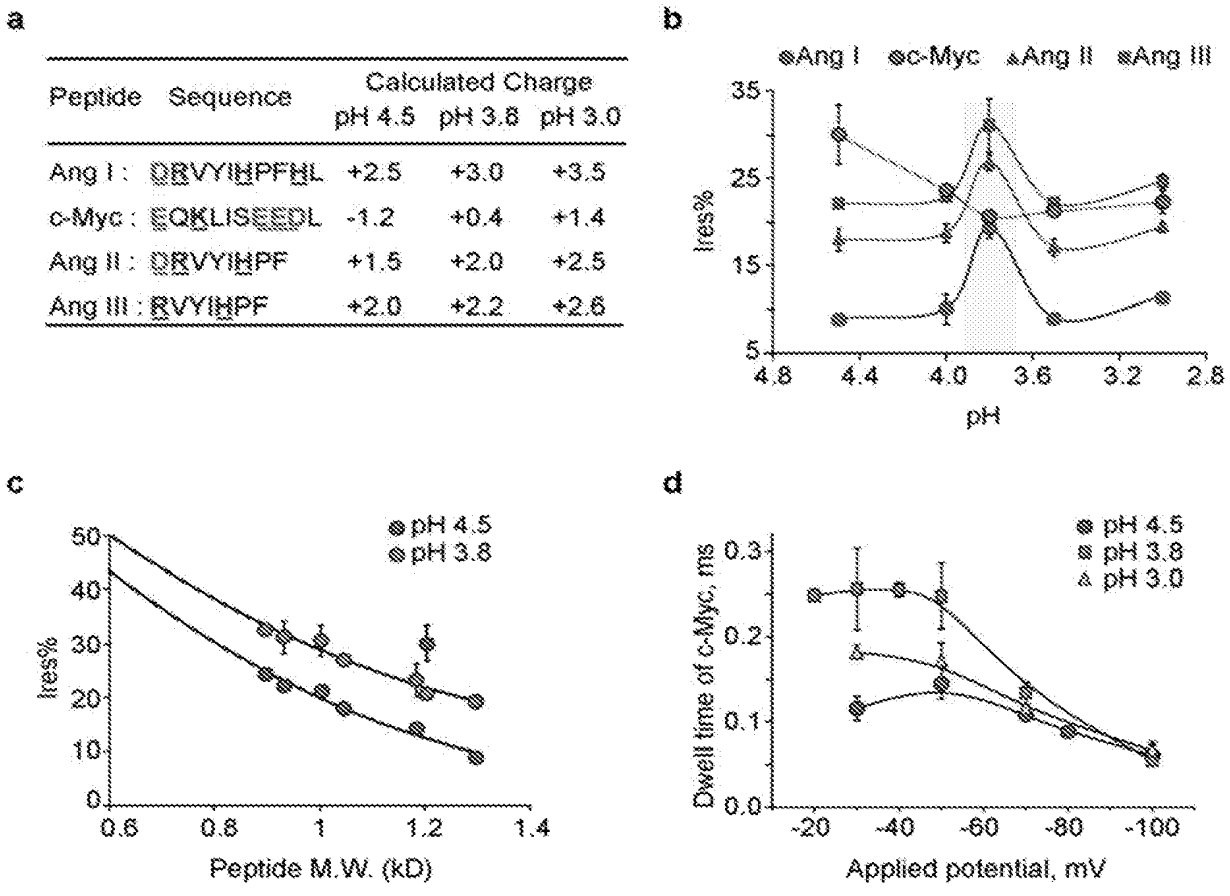
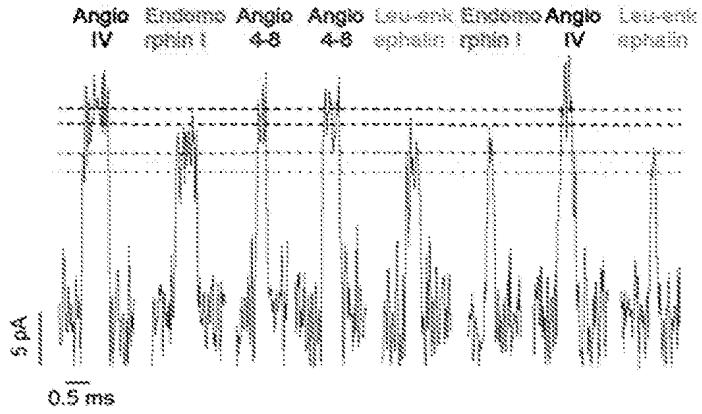


Figure 6

a

Peptide sequence		M.W (g/mol)	Ires% (-50 mV)
Angio IV:	V Y I H P F	774.4	1.1 ± 0.8
Angio 4-8:	Y I H P F	675.8	8.2 ± 0.4
Endomorphin I:	Y P W F	610.7	19.2 ± 0.5
Leu-enkephalin:	Y G G F L	555.6	34.5 ± 2.4

b



c

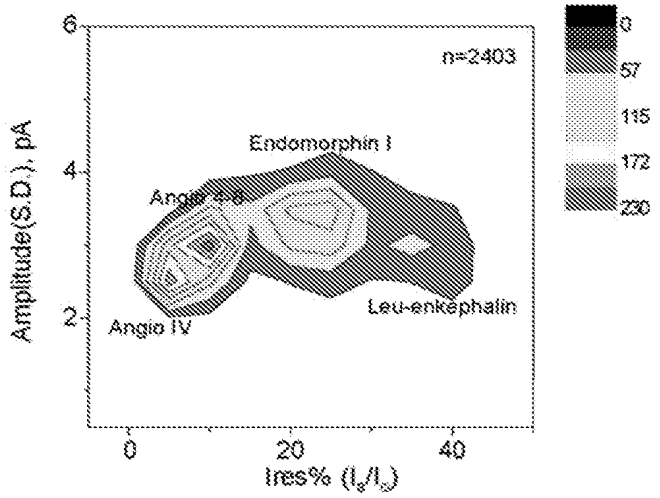
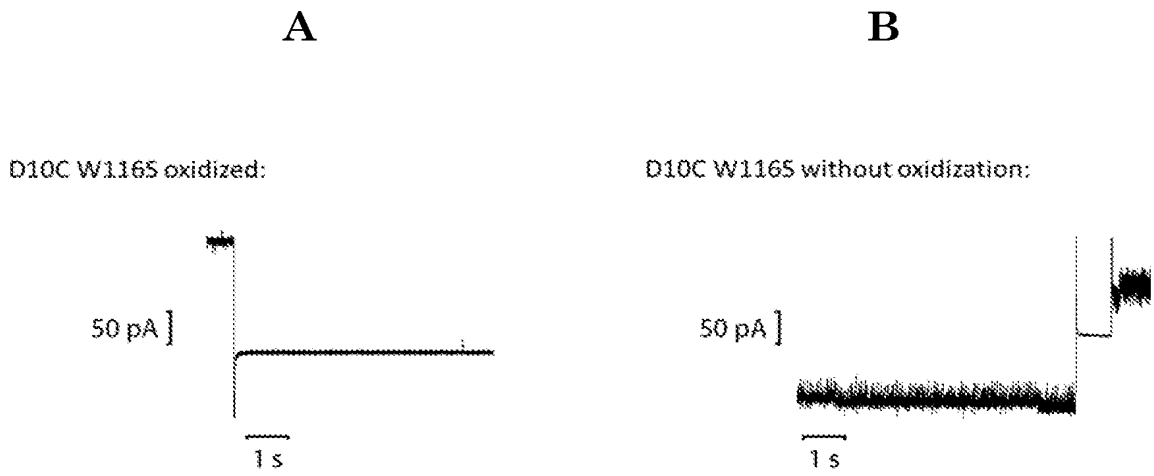


Figure 7


A
B
D10C W1165 oxidized:
50 pA
1 s
D10C W1165 without oxidization:
50 pA
1 s

INTERNATIONAL SEARCH REPORT

International application No
PCT/NL2019/050588

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/435 G01N33/487
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, CHEM ABS Data, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2018/012963 A1 (UNIV GRONINGEN [NL]) 18 January 2018 (2018-01-18)	12,19-21
A	abstract page 21, line 19 - page 23, line 5; claims 1-41; table 3	1-11, 13-18

Y	JASON MACRANDER ET AL: "Evolution of the Cytolytic Pore-Forming Proteins (Actinoporins) in Sea Anemones", TOXINS, vol. 8, no. 12, 8 December 2016 (2016-12-08), page 1-16, XP055556874, DOI: 10.3390/toxins8120368	12,19-21
A	abstract page 9, paragraph 1	1-11, 13-18

	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 December 2019	17/12/2019

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Gurdjian, Didier

3

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE UniProt [Online] 1 September 2009 (2009-09-01), "RecName: Full=DELTA-actitoxin-Afr1a {ECO:0000303¦PubMed:22683676}; Short=DELTA-AITX-Afr1a {ECO:0000303¦PubMed:22683676}; AltName: Full=Alpha-helical pore-forming toxin {ECO:0000303¦PubMed:21300287, ECO:0000303¦PubMed:22728830, ECO:0000303¦PubMed:25716479}; Short=PFT {ECO:0000303¦PubMed:21300287, ECO", XP002796191, retrieved from EBI accession no. UNIPROT:B9W5G6 Database accession no. B9W5G6 cited in the application the whole document -----	12,19-21
Y	BELLOMIO A ET AL: "Purification, cloning and characterization of fragaceatoxin C, a novel actinoporin from the sea anemone Actinia fragacea", TOXICON, ELMSFORD, NY, US, vol. 54, no. 6, 1 November 2009 (2009-11-01), pages 869-880, XP026499390, ISSN: 0041-0101, DOI: 10.1016/J.TOXICON.2009.06.022 [retrieved on 2009-06-27] abstract page 874, right-hand column, paragraph 1; figures 2,3 -----	12,19-21
A	ARIEL E MECHALY ET AL: "Structural Insights into the Oligomerization and Architecture of Eukaryotic Membrane Pore-Forming Toxins", STRUCTURE, ELSEVIER, AMSTERDAM, NL, vol. 19, no. 2, 6 November 2010 (2010-11-06), pages 181-191, XP028357696, ISSN: 0969-2126, DOI: 10.1016/J.STR.2010.11.013 [retrieved on 2010-12-21] abstract; figure 2 ----- -/--	1-5,12, 13

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WLOKA CARSTEN ET AL: "Alpha-Helical Fragaceatoxin C Nanopore Engineered for Double-Stranded and Single-Stranded Nucleic Acid Analysis.", 26 September 2016 (2016-09-26), ANGEWANDTE CHEMIE (INTERNATIONAL ED. IN ENGLISH) 26 09 2016, VOL. 55, NR. 40, PAGE(S) 12494 - 12498, XP002788982, ISSN: 1521-3773 abstract; figure 1	1-4,12, 13
A	SIMONE FURINI ET AL: "Model-Based Prediction of the [alpha]-Hemolysin Structure in the Hexameric State", BIOPHYSICAL JOURNAL, vol. 95, no. 5, 1 September 2008 (2008-09-01), pages 2265-2274, XP055556977, AMSTERDAM, NL ISSN: 0006-3495, DOI: 10.1529/biophysj.107.127019 abstract; figure 1	1-4,12, 13

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/NL2019/050588

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2018012963 A1	18-01-2018	AU 2017295442 A1	07-02-2019
		BR 112019000680 A2	24-04-2019
		CA 3030704 A1	18-01-2018
		CN 109890980 A	14-06-2019
		EP 3485029 A1	22-05-2019
		JP 2019533637 A	21-11-2019
		US 2019292235 A1	26-09-2019
		WO 2018012963 A1	18-01-2018